DOCUMENT RESUME

ED 216 185 CE 032 505

AUTHOR Freeberg, Norman E.; Rock, Donald A.
TITLE What to Do with What We Know and What to Do about What We Don't Know.
PUB DATE Mar 82
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Databases; Data Collection; Educational Policy; *Educational Research; *Education Work Relationship; *Employment Programs; Evaluation Criteria; Evaluation Methods; Followup Studies; Information Needs; Information Utilization; Longitudinal Studies; Outcomes of Education; Program Effectiveness; *Program Evaluation; *Research Design; Research Methodology; *Research Needs; Research Problems; Summative Evaluation; Surveys; Youth; Youth Employment
IDENTIFIERS Standard Assessment System; *Youth Career Development Program; Youth Employment and Demonstration Projects Act

ABSTRACT

Because of various flaws in their design, most research studies that have up to now examined the transition of youth to the work setting have failed to pinpoint ways of improving existing programs. Many of these design flaws (including lack of a uniform set of outcome measures, failure to routinely incorporate control groups, and failure to design for replication) have been overcome in a study that examined the effectiveness of the Youth Career Development (YCD) Program. The study's design involved the use of a Standard Assessment System (SAS) that could serve as a core system with a range of behavioral constructs that would be applicable to program objectives across the variety of Youth Education and Demonstration Projects Act (YEDPA) programs. Included in the SAS were a battery of measures of vocationally oriented attitudes and knowledge, performance outcome survey instruments, and surveys to collect demographic data. Based on a summative approach, the study focused on determining what works best for whom when assessing particular programs. After analyzing the results of the survey as well as the design of the study itself, researchers found that both the YCD and the measurement system used to evaluate it were extremely effective. Recommendations called for further research to determine why some programs work and others do not. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED216185

WHAT TO DO WITH WHAT WE KNOW AND WHAT TO DO ABOUT WHAT WE DON'T KNOW

Norman E. Freeberg
Donald A. Rock
Educational Testing Service

AERA Symposium on the Transition of Youth from Education to Work: Issues and Answers

New York, March 21, 1982

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PART I: DEVELOPMENT OF THE KNOWLEDGE BASE

The initial issues to be considered prior to discussing what we know of the transition of youth to the work setting is: "How did we get to know what we know?" and the nature and quality of information from which that knowledge is derived.

1. Purpose of the Youth Employment Demonstration Projects ACT (YEDPA)

Enacted by Congress in 1977, YEDPA was intended to create special demonstration projects to explore the feasibility and effects of alternative and innovative approaches for dealing with barriers to youth unemployment. Information from more than a decade of previous studies had failed to pinpoint ways of improving on what had been undertaken because of such weaknesses in assessing alternative approaches as: (a) lack of a uniform set of within program and post-program outcome measures obtained in a longitudinal time frame, (b) failure to incorporate comparison or control groups routinely for contrasting program effects and (c) failure to design for replication in numerous localities with differing site characteristics.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Norman E. Freeberg

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED232505

The techniques and study approaches for building a knowledge base that overcomes these deficiencies, are outlined in this paper and the results and utility of that knowledge are described for one large-scale prototypical YEDPA program--The Youth Career Development (YCD) Program. YCD was conducted with a participant sample of some 1700 disadvantaged (CETA qualified) high school seniors at 30 sites throughout the country; its overall aim being to provide career development services and skills that would enhance student success in the transition from school to work.

2. The Evaluation Measurement System

A critical element in determining the effectiveness of any youth program is the quality of the measurement information obtained for evaluation purposes. A set of measures was chosen that could serve as a "core" system with a range of behavioral constructs that would be applicable to program objectives across the variety of YEDPA experimental programs. In addition, it was necessary that such measures be appropriate in design and content for economically and culturally disadvantaged (largely minority group) adolescents. These were chosen from the best existing "off-the-shelf" measures, with the intent of examining further their value for use with the YEDPA population.

A Standard Assessment System (SAS) was assembled that contained three major segments for assessing trainee performance. The first is a battery of 7 measures of vocationally-oriented attitudes and knowledge consisting of scales designated as Vocational Attitude, Job Seeking Skills, Job Holding Skills,

Self Esteem, Job Knowledge, Work Related Attitudes, and Sex Stereotyping of Occupations. The second segment is comprised of three performance outcome survey instruments. One, for use at program completion,measures trainee "success" at the time of leaving the program in terms of dimensions of work motivation, training program adjustment, social adjustment, vocational expectations and planning competency. Another is for use at various post-program follow-up time periods (e.g., 3 and 8 month periods in our studies to date) and contains items that fall under dimensions of Social and Vocational Adjustment, Job Success and Satisfaction, Job Search Motivation, and Job Planning Capability. The third segment of the SAS contains a variety of demographic information (age, sex, race; economic, educational, and labor force status, etc.) and a short test of verbal ability--all of which are intended for use in describing sample composition and as adjustment or "equating" variables for covariate analytical purposes.

3. Study Design and Data Collection

Design of this longitudinal study of a youth career development program (and of most of the other YEDPA demonstrations incorporated in our data base using the Standard Assessment System) was of a quasi-experimental type. This entailed the pretesting of a participant (treatment) group with the 7 measures of the assessment battery at the time of entry into the program and posttesting them with the same set of measures following program participation. Contrasts in score change to define the significance of gain were

made possible by using a comparable or control group of students tested at about the same time periods, but for whom there was no intervening program participation.

Follow-up survey information was obtained at 3 and 8 months after program completion (i.e., following completion of the academic year) for those students who could be located and would voluntarily respond. Potential problems in data collection and methodology were introduced by the fact that the agencies responsible for the conduct of the programs were solely responsible for the evaluation data collection as well. High turnover rates at a number of project sites for personnel engaged in that effort made it difficult to pass on established data collection procedures. Replacement of trainee dropouts resulted in rolling admissions and rolling terminations with consequent breaks in the normal data collection sequence and differential pretest to posttest intervals (i.e., in length of time of program enrollment). In addition,no random selection or random assignment to participant and control samples was feasible,in a voluntary program so that possible sample biases could be introduced from the interaction of student self selection and training (treatment) effects.

4. Analytical Approach

The analytical approach,to which the design and the available data lend themselves,is essentially a summative one concerned with determining what works best and for whom in assessing the effectiveness of a particular program. (The causal question of why

is the one with broader policy implications, that can be dealt with only across a number of programs of similar type. Techniques for applying structural analysis to that question will be mentioned in Part II of this paper.)

The primary questions to be dealt with in this summative analysis involve:

(1) To what extent do within program changes occur and for which of the tested behavioral constructs? This entails use of the method of Analysis of Covariance with adjustments applied in gain score analyses for pre-existing differences between participant and control groups (i.e., covariate adjustments based on variables of reading level, sex, family income level, advantaged/disadvantaged status, ethnic group membership, and previous employment).

(2) Are gains in the constructs measured related to later vocational performance outcome status? (Based on a residualized gain score analysis that results in part correlations with control for initial or pretest status only.)

(3) Is there a differential impact of the program on differing subgroups of participants in their within program gain or is there generally uniform gain across all groups regardless of differing background characteristics? (The ANCOVA type of computation used for this purpose yields demographic correlates of gain which are independent of the participant's pretest score.)

(4) Was there significant benefit to participants in their post-program performance, during school-to-work transition, 3 and 8 months after leaving the program? (Contrasts in participant and control group means for follow-up survey variables were obtained from ANCOVA with adjustments for initial student status--i.e., reading level and the same demographic variables as previously applied.)

(5) What was the extent of bias in the longitudinal samples obtained from pretest through 8 month follow-up that may have resulted from selective attrition? (This is dealt with by examination of sample composition for participants and controls on distributional patterns for key demographic variables.)

(6) How well did the measurement properties of the assessment system hold up? (As determined from scale internal characteristics as well reliabilities and predictive validities.)

PART II: WHAT DID WE FIND OUT? HOW CAN IT BE USED AND WHAT CAN WE DO ABOUT LEARNING MORE?

From the information contained in the knowledge base and the analytical approach outlined above, it was found that:

(1) The program participants, explosed to a career development program, showed improvement over the course of training as reflected in gains on all 7 measures of the test battery with statistically significant gains on 5 of the 7 (i.e., vocational attitudes, job holding and seeking skills, and sex stereotyping of occupations).

(2) The extent of the change on all 7 of the measures was positively related to one or more important career-oriented outcomes;particularly the status level of the full-time jobs obtained by former participants at 3 and 8 month post program periods (i.e., those who gained more on the tested constructs obtained higher status jobs). Gains on most of those 7 measures were also related to whether or not the individual was doing something "useful" after the program (i.e., working full or part time and/or going to school full or part time as opposed to "doing nothing"). There was, however, uniformly no relationship of gain to whether or not the participant had achieved full-time employment.

(3) Differential program impact, in terms of test score gains for any particular subgroup of participants was small and scattered indicating that, for the most part, favorable effects were achieved "across the board"--except for one participant inital status characteristic and that was reading ability. Gain achieved on all 7 measures over the course of the program was obtained for those of higher reading skill.

(4) Favorable post-program impact on a variety of job and social variables was found for participants at the 3 and 8 month outcome periods: including higher status jobs (for those who obtained full time employment), a greater likelihood of obtaining full time employment (amont those who did not go on to post-secondary school) and a greater likelihood of engaging in some form of "useful activity" (school and/or work full or part time).

(5) Examination of possible sample bias, based on whether attrition was distributed uniformly throughout the sample subpopulation, indicated that the representativeness of the sample at any longitudinal time period held up reasonably well. Where shifts in distributional patterns were found (such as a general increase in females of about 5% by 8 month follow up), the shifts occurred in much the same way for participant and control groups. Thus, the important participant-control comparisons were not likely to be distorted and any results could be generalized reasonably to the youth population that originally entered the program.

(6) The measurement system held up particularly well. Internal characteristics of the 7 scales of the assessment battery (means, variances, score ranges) were highly similar for participant and control groups and for previous samples on which the measures were developed. The same is true of the reliability coefficients which were of sufficient magnitudes for the intended evaluation purposes (r's range from mid 50's to .90).

Predictive validities of the 7 measures of the battery, using 3 and 8 month outcome variables as criteria, are numerous though of very modest levels (significant r's from the mid teens to mid .20'--with the Work Related Attitudes scale the one with the best validities overall). A degree of criterion value (relevance) was found for the follow up survey instruments, not only in the predictability of their outcome variables by the test constructs, but in the pattern of relationships between those outcome variables at the 3 and 8 month follow up periods.

The major implications for applying these findings to youth training programs (beyond the obvious one that there are benefits to be derived from exposing students to a career development program) are: (1) the types of constructs measured that were intended to reflect program objectives (i.e., job knowledge, job seeking skills, job holding skills, attitudes toward work, self confidence) are worth incorporating in training curricula, since changes (gains) in those aspects of vocational attitudes and knowledge are related to subsequent career-oriented performance outcomes; (2) since the trainee subgroup possessing superior reading ability tends to be one that gains (benefits) more during the course of program participation, a possible role for verbal skill remediation incorporated early in the program is implied (3) the components of the present measurement system are worth applying in the evaluation of the YEDPA type of youth training population, based on the measurement properties demonstrated.

Where Do We Go From Here?

If the summative analysis based on the total data base shows that some programs "work" while others show little or no gain, when compared to non-program youth, then certain policy questions need to be addressed. The policy questions are:

(1) Which labor market outcomes were most successfully achieved through program participation?

(2) Why will certain types of programs work better than others?

(3) When did they work? That is, was the program impact felt immediately following participation or did it develop over time and/or did it start strong and then attenuate over time?

(4) What kinds of individuals seem to profit most from which type of program participation?

(5) What are the cost benefit ratios achieved for those programs which seem to work?

Although the above questions are stated as if they are mutually exclusive, in reality they are not independent. For example, one might expect on further analysis to find that some programs work better with respect to certain types of youth and with respect to certain kinds of goals. In short, one might expect to see interactions between types of youth, labor market outcomes, and program types.

What kind of data and analysis are required to answer the above questions? Three basic forms of information are needed. First, information on the individual's abilities, attitudes, work experiences, and demographic characteristics; secondly, program process and descriptive information; and finally, the policy (causal) question requires labor market outcome data gathered at selected time periods. The present data bank has the information needed to address these policy questions. In addition to the individual descriptive data already observed, we have collected process and descriptive information at the program site level. This information allows one to classify program sites with respect to similarity of processes, type of delivery systems and primary goals.

Figure 1 of your handout presents a classification of program sites by type of input (in-school vs. out-of-school youth) and their primary objectives. For example, pre-employment job search serves out-of-school youth and attempts primarily to place them in jobs and not in schools. School-to-work serves in-school youth and attempts to either place trainees in jobs or in school. The "double checks" indicate greater emphasis on a particular outcome. Programs are being compared within categories, in order to estimate the variability of program effects within those categories. It may well be that the variability

within program categories, with respect to achieving proposed outcomes, is sufficiently large compared to the between category variance that one may conclude that there may be "exemplary" programs within each category. It then becomes a matter of finding out the "why," "when," and at "what cost" the gains are achieved within the various program categories.

The relationship between program impact and process information can then be analyzed in path analytic models which incorporate individual and contextual information in an attempt to explain why certain programs have a more positive impact than others. In this particular analysis, the term "contextual" implies a broad range of descriptions gathered from the process questionnaire. Contextual scores will be assigned to individuals at the program site level and will include such program characteristics as: public and private sector linkages, staffing characteristics, facilities, size, and types of service delivery systems. Information will also be analyzed with respect to allocation of effort between program components such as occupational training, placement, counseling, and follow-up or basic skills services. Detailed cost information will be used to estimate cost benefit ratios.

The contextual analysis will be designed not only to evaluate the relative impact of individual characteristics and programmatic variables on program outcomes, but also to estimate the interaction of these effects. It is hoped that one of the products of this analysis of the total data base will be a "blueprint" or hypothetical profile of the most "cost effective" program(s). That is, one should be able to put together a listing of characteristics, both individual and programmatic, which define an exemplary program(s). It would be anticipated that such a blueprint would prove helpful in the program design, for any new private or public initiative with respect to youth employment training programs.

Figure 1

Categories of Programs by Total Data Base

Primary Goals

	Employment	Employment Training	School
School to Work (in school)	✓✓	✓	✓✓
Summer (in school)	✓	✓	✓✓
Summer (out of school)	✓✓	✓	✓
Pre-employment job search (out of school)	✓✓		
Training and apprenticeship (out of school)	✓✓	✓✓	
Work Experience (out of school)	✓✓	✓✓	
School to School (in school)			✓✓